Heterogeneous Treatment Effect Analysis

Ben Jann

ETH Zurich

In cooperation with Jennie E. Brand (UCLA) and Yu Xie (University of Michigan)

German Stata Users Group Meeting

Berlin, June 25, 2010

Introduction

- Methods for causal inference from observational data have received much attention in the last two decades or so, especially in econometrics, but also in many other fields.
- Starting point of this literature is the Rubin Causal Model (a.k.a. Potential Outcomes Model a.k.a Counterfactual Causality).
- Assume a binary treatment variable D and let Y^1 and Y^0 be the potential outcomes with and without treatment, respectively. The treatment effect for individual i is then simply the difference between the potential outcomes, that is

$$\delta_i = Y_i^1 - Y_i^0$$

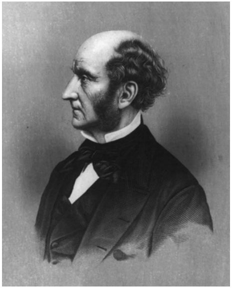
- The fundamental problem of causal inference, however, is that we can only observe Y_i^1 *or* Y_i^0. One of the potential outcomes must be counterfactual because what we observe is

$$Y_i = \begin{cases} Y_i^1 & \text{if } D_i = 1 \\ Y_i^0 & \text{if } D_i = 0 \end{cases}$$

Introduction

- The idea of defining causality in terms of potential outcomes is not new:

> *Thus, if a person eats of a particular dish, and dies in consequence, that is, would not have died if he had not eaten of it, people would be apt to say that eating of that dish was the cause of his death.*[1]



John Stuart Mill (1806–1873)

[1] John Stuart Mill (2002). A System of Logic. Reprinted from the 1981 edition (first published 1843). Honolulu, Hawaii: University Press of the Pacific. P. 214.

Introduction

- A basic paradigm of the literature based on the potential outcomes model is that there can be individual heterogeneity in treatment effects, which stands in contrast to traditional regression modeling assuming constant parameters.
- The view that treatment effects can be heterogeneous led to new methods for causal inference and also to new uses and interpretations of existing methods (e.g. LATE interpretation of IV estimators, revival of matching and regression discontinuity designs).
- Surprisingly, however, not much attention is usually paid to the explicit analysis of the heterogeneity of treatment effects in applied studies.
- The basic quantity of interest is the average treatment effect (ATE)

$$ATE = E[\delta_i] = E[Y_i^1 - Y_i^0] = E[Y_i^1] - E[Y_i^0]$$

  or sometimes the average treatment effect on the treated ($ATT = E[\delta_i | D_i = 1]$) or the average treatment effect on the untreated ($ATC = E[\delta_i | D_i = 0]$).

Introduction

- Why should we care about analyzing heterogeneous treatment effects?
- The naive estimator of the average treatment effect based on observational data can be decomposed as

$$
\begin{aligned}
NATE &= E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 0] \\
&= E[\delta_i] + \underbrace{E[Y_i^0|D_i = 1] - E[Y_i^0|D_i = 0]}_{\text{pre-treatment heterogeneity bias}} \\
&\quad + (1 - E[D_i])\underbrace{(E[\delta_i|D_i = 1] - E[\delta_i|D_i = 0])}_{\text{treatment-effect heterogeneity bias}}
\end{aligned}
$$

- The focus of most estimation approaches is to eliminate the first type of bias, but also the second type of bias might threaten the validity of causal inference.

Introduction

- For example, in the literature on economic returns to higher education various theories have been proposed that imply heterogeneous effects depending on the probability to go to college.
  - Human-capital theory in economics predicts *positive selection* into treatment, because people choose to go to college based on the expected economic returns. This is a widely accepted view.
  - More sociologically oriented literature suggests that college attendance is strongly influenced by social origin, which leads to *negative selection* into treatment under certain conditions.
- To evaluate these theories it is therefore crucial to analyze how treatment effects vary with treatment probability.
- Ultimately, believes about the mechanisms at play determine educational policy.

Analysis of Heterogeneous Treatment Effects

- To support the analysis of treatment-effect heterogeneity we developed a new tool called `hte`.
- The approach of `hte` is to assume, at least provisionally, conditional unconfoundedness given a set of covariates and use propensity score stratification to estimate treatment effects at various points over the range of the propensity score.
- These strata-specific effects are then analyzed to determine whether there is a trend in treatment effects.

Algorithm

- The `hte` algorithm consists of four basic steps.
  1. Estimation of the propensity score (i.e. the conditional probability to receive treatment).
     - ⋆ `hte` uses `probit` or `logit`, but it is also possible to manually estimate the propensity score beforehand and then provide it to `hte`.
  2. Construction of balanced propensity score strata.
     - ⋆ `hte` calls the `pscore` command for this purpose.[1]
  3. Estimation of strata-specific average treatment effects.
     - ⋆ In each stratum, a regression model on treatment is estimated, optionally including control variables to account for remaining covariate imbalance within strata.
  4. Estimation of the trend of treatment effects across propensity score strata.
     - ⋆ `hte` regresses the strata-specific treatment effects on strata rank using variance weighted least squares (`vwls`; with the variance based on the standard errors of the strata specific treatment effects).

[1] Becker, S. O., A. Ichino (2002). Estimation of average treatment effects based on propensity scores. *The Stata Journal* 2:358–377.

Example

- Syntax
- Example
- An application of the procedure can be found, for example, in:
  - Brand, J. E., Y. Xie (2010). Who Benefits Most From College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education. *American Sociological Review* 75:273–302.

Syntax

Viewer (#2) [help hte]

help hte

Advice | Contents | What's New | News

help hte

Title

hte — Heterogeneous Treatment Effect Analysis

Syntax

hte *depvar1* [*depvar2* ... =] *treatvar* *indepvars* [*if*] [*in*] [*weight*] [, *options*]

hte graph [, **level(#)** *graph_options*]

options	description
Main	
alpha(#)	set significance level for balancing tests
pscore_options	options as described in help **pscore** (except **level()**)
join(*list***)**	merge specified strata
autojoin[(#)]	merge small strata at low end or high end
by(*groupvar***)**	repeat analyses for groups defined by *groupvar*
separate	construct propensity score strata separately for each by-group
controls(*clist***)**	control variables for within-strata models
estcom(*command***)**	set estimation command for within strata models; default is **regress**
estopts(*options***)**	options to be applied to within-strata models
noisily	display output from **pscore** and individual models
level(#)	set confidence level; default is **level(95)**
listwise	use listwise deletion to handle missing values
casewise	synonym for **listwise**
Graph	
nograph	suppress graph
outcomes(*numlist***)**	display results for specified outcomes
marker_options	change look of markers
lineopts(*options***)**	change look of fitted lines
ciopts(*options***)**	change look of confidence intervals

Example

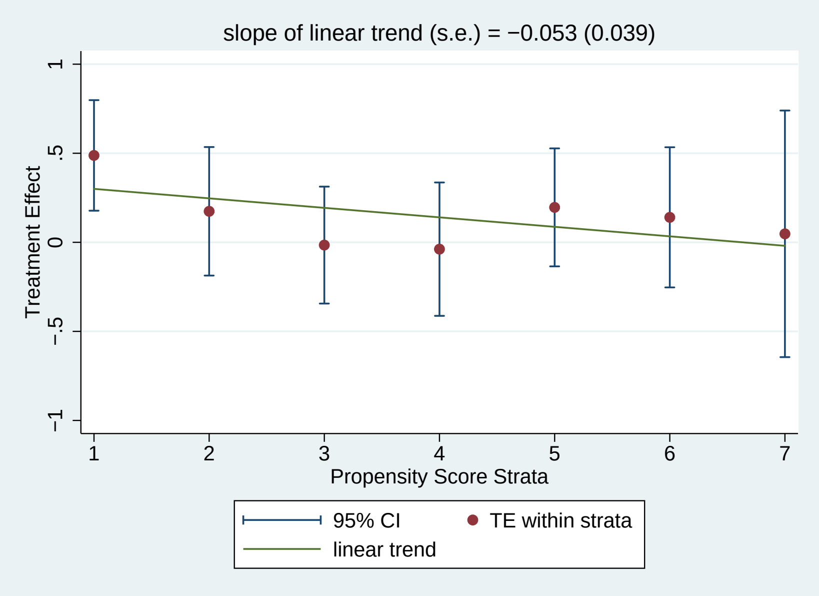
```
. hte highschool childcare                                      ///
>     female peduclow peduchigh lnhhinc motherlfp immigrant     ///
>     siblings1-siblings3 cohort1991-cohort1995 east
```

Number of obs = 594

highschool	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
TE by strata						
1	.4878277	.158317	3.08	0.002	.1775321	.7981234
2	.1740196	.1840194	0.95	0.344	-.1866517	.5346909
3	-.0155844	.1674786	-0.09	0.926	-.3438365	.3126677
4	-.0384615	.1910365	-0.20	0.840	-.4128862	.3359632
5	.1960784	.1689491	1.16	0.246	-.1350557	.5272126
6	.1401515	.2007209	0.70	0.485	-.2532543	.5335573
7	.047619	.3531523	0.13	0.893	-.6445468	.7397849
Linear trend						
_slope	-.0532744	.0388194	-1.37	0.170	-.1293591	.0228102
_cons	.3533125	.1521936	2.32	0.020	.0550185	.6516065

TE = treatment effect

Example



slope of linear trend (s.e.) = −0.053 (0.039)
Treatment Effect
1
.5
0
−.5
−1
1
2
3
4
5
6
7
Propensity Score Strata
95% CI
TE within strata
linear trend

Work in Progress

- Plans for `hte`:
  - Optional nonparametric estimation of propensity score.
  - Improve the balanced propensity score strata algorithm and provide better output (descriptive information on strata, balancing tests, etc.).
    - Requires a rewrite of `pscore`.
  - Automate within strata covariate adjustment.
  - Formal tests for treatment-effect heterogeneity.
  - Improve level-2 estimation.
- `hte2`: fully nonparametric approach
  - Estimate observation-specific counterfactual outcomes.
  - Use non-parametric estimators to analyze the trend in treatment effects over propensity score or across the values of covariates.
  - Example.

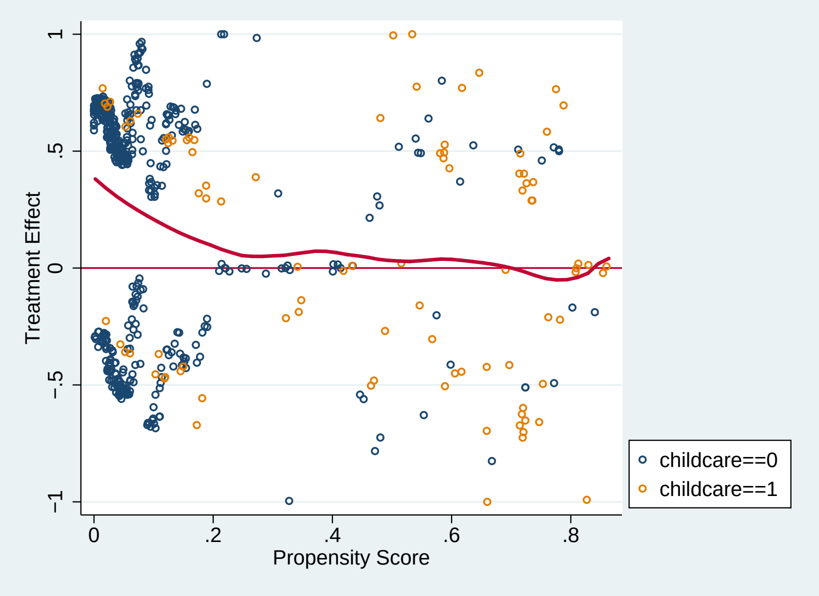
Example

```
psmatch2 childcare                                              ///
    female peduclow peduchigh lnhhinc motherlfp immigrant       ///
    siblings birthyr east,                                      ///
    outcome(highschool) kernel bw(0.025) ate

gen double treatefct = cond(childcare==0,                       ///
    _highschool-highschool, highschool-_highschool)

twoway scatter treatefct _pscore if childcare==0 & _support==1, ///
            jitter(2) msym(oh)                                  ///
    || scatter treatefct _pscore if childcare==1 & _support==1, ///
            jitter(2) msym(oh) psty(p4)                         ///
    || lpoly   treatefct _pscore if  _support==1,               ///
            pstyle(p6) lw(*2) degree(1)                         ///
    yline(0) xti(Propensity Score) yti(Treatment Effect)        ///
    legend(order(1 "childcare==0" 2 "childcare==1")             ///
            cols(1) position(4))
```

Example



Treatment Effect
1
.5
0
−.5
−1
0
.2
.4
.6
.8
Propensity Score
childcare==0
childcare==1

Thank you for listening!